

**Department of Pediatrics**

Section of Endocrinology

13123 East 16th Avenue

Box 265

Aurora, CO 80045

Office: 720-724-4358

deborah.glueck@ucdenver.edu

March 26, 2018

Dear Professors Balakrishnan, Schucany, Kshirsagar and Gentle:

I submit for your consideration the manuscript, "Power for balanced linear mixed models with complex missing data processes," for publication in *Communications in Statistics-Theory and Methods*.

The manuscript presents a novel non-central power approximation for the Wald test in a wide class of linear mixed models. The approximation adjusts for complex missing data processes.

We recently submitted the manuscript to *Statistical Methods in Medical Research*. The reviewers pointed out that there was an existing GEE approach to approximate power in study designs with clustered data that seemed to obviate the need to improve upon mixed model approaches. We revised the manuscript to include citations to support our position that GEE methods can have high Type I error inflation in studies with relatively small sample sizes (). The inflation makes the GEE methods unsuitable for the class of studies we consider in our work. Missing data compounds the issue by effectively reducing the sample size even further.

Given the focus of the paper is on missing data power approximation, we chose not include any simulations to support our case for the study scenarios of interest. However, we did conduct a small simulation study to provide first-hand evidence of the GEE method's unfavorable performance for the study designs we consider. We compared the GEE empirical Type I error rates to those for the mixed model Wald test for studies with different sample sizes and amounts of missing data. The simulation study demonstrates that fitting GEE models to correlated data with small sample sizes and missing data produces inflated Type I error rates. The inflation increases with decreasing effective sample size, *i.e.*, the expected sample size after accounting for the missing data process. By contrast, the mixed model Wald test has well-controlled Type I error rates for most experimental conditions and has considerably smaller Type I error rate inflation than the GEE method. While we do not believe the simulation study is within the scope of our manuscript, we provide the results here, in the cover letter, as an aid in assessing the impact of our work. A more detailed description of the simulation study methods and results are included in the Attachment.

Each author contributed substantially to the manuscript. None of the authors have any conflicts of interest to declare. The mathematical derivations and simulation studies were funded by NIH R01GM121081 (Glueck, Dabelea, Muller), NIH R25GM111901 (Glueck, Muller) and NIH G13LM011879 (Glueck, Muller). The background research for the illustrative examples was funded by NIH R01DK076645 (Dabelea), NIH UH3OD023248 (Dabelea), AHA16MCPRP29710005 (Sauder), NIH R01 HD043770 (Schenkman), CCTSI TL1RR025778 (Schenkman), NIH P30 DK048520 (Hill), K23 NS052487 (Hall), and the Parkinson's Disease Foundation.

Thank you for your consideration of the manuscript. Please feel free to contact me if you require any further assistance. I look forward to hearing from you

Sincerely,

Deborah H. Glueck, Ph.D.

Visiting Professor

Department of Pediatrics

University of Colorado School of Medicine

12474 E. 19th Avenue

Building 402, Room 219

Mail Stop F426

Aurora, CO 80045

**ATTACHMENT**

The simulation study experimental conditions represent a subset of conditions we present in Scenarios 4 and 5 within our manuscript. The study design for Scenario 4 had four exposure groups and three repeated measurements over time. The study design for Scenario 5 had two exposure groups and five repeated measurements over time. In both scenarios, we considered a test of a time-by-treatment interaction. The nominal Type I error rate (α-level) for Scenario 4 was 0.01 and for Scenario 5 was 0.05. The planned sample sizes included and for Scenarios 4 and 5, respectively. For each planned sample size, we considered a range of missing data patterns and amounts of missing data, resulting in different expected numbers of observed data points, or different *effective sample sizes*.

We replicated each experimental condition 10,000 times. For each replication, we attempted to fit a GEE model and a mixed model. All modeling was completed in SAS 9.4. For the GEE models, we used the PROC GENMOD procedure with an unstructured correlation. Mixed models were fit using the PROC GLIMMIX procedure with an unstructured covariance and Kenward-Roger denominator degrees of freedom. For each GEE model that converged, we used the GEE estimates to calculate a Wald statistic and compute a p-value for a central chi-square reference distribution with either 12 (Scenario 4) or 10 (Scenario 5) degrees of freedom. For each mixed model that converged, we calculated the mixed model Wald statistic and computed p-values for an reference distribution with either 12 (Scenario 4) or 10 (Scenario 5) numerator degrees of freedom and Kenward-Roger denominator degrees of freedom. Empirical power was computed as the number of p-values less than or equal to α divided by the number of replicates for which the model converged.

Tables 1 and 2 show that, for small sample sizes, either planned or effective, the Type I error rate for the GEE method is highly inflated. As the planned and effective sample sizes increase, the Type I error for the GEE method stabilizes, though it never achieves the nominal rate. The Type I error results for the mixed model approach achieve the nominal rate for all tests experimental conditions for Scenario 4. For Scenario 5, the mixed model Type I error rates decrease with increasing planned and effective samples sizes, similar to the GEE method results. However, the mixed model empirical Type I error rates are not as highly inflated as those for the GEE method and achieve the nominal Type I error rate as the sample sizes increase. We conclude that a mixed model approach is more appropriate than a GEE approach for studies with small planned sample sizes or for studies that are expected to have many missing measurements.

Table 1. Mixed model and GEE empirical Type I error rates for Scenario 4 and an α-level of 0.01.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Planned Sample Size | Effective Sample Size | Nominal  Type I Error Rate | GEE  Empirical  Type I Error Rate | Mixed Model Empirical  Type I Error Rate |
| 40 | 32 | 0.01 | 0.22 | 0.01 |
| 40 | 33 | 0.01 | 0.18 | 0.01 |
| 40 | 35 | 0.01 | 0.16 | 0.01 |
| 40 | 36 | 0.01 | 0.13 | 0.01 |
| 40 | 40 | 0.01 | 0.10 | 0.01 |
| 80 | 64 | 0.01 | 0.06 | 0.01 |
| 80 | 67 | 0.01 | 0.06 | 0.01 |
| 80 | 69 | 0.01 | 0.05 | 0.01 |
| 80 | 72 | 0.01 | 0.05 | 0.01 |
| 80 | 80 | 0.01 | 0.04 | 0.01 |
| 160 | 128 | 0.01 | 0.03 | 0.01 |
| 160 | 133 | 0.01 | 0.02 | 0.01 |
| 160 | 139 | 0.01 | 0.02 | 0.01 |
| 160 | 144 | 0.01 | 0.02 | 0.01 |
| 160 | 160 | 0.01 | 0.02 | 0.01 |
| 320 | 256 | 0.01 | 0.02 | 0.01 |
| 320 | 267 | 0.01 | 0.02 | 0.01 |
| 320 | 277 | 0.01 | 0.02 | 0.01 |
| 320 | 288 | 0.01 | 0.02 | 0.01 |
| 320 | 320 | 0.01 | 0.01 | 0.01 |
| 400 | 320 | 0.01 | 0.02 | 0.01 |
| 400 | 333 | 0.01 | 0.01 | 0.01 |
| 400 | 347 | 0.01 | 0.02 | 0.01 |
| 400 | 360 | 0.01 | 0.01 | 0.01 |
| 400 | 400 | 0.01 | 0.01 | 0.01 |

Table 2. Mixed model and GEE empirical Type I error rates for Scenario 5 and an α-level of 0.05.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Planned Sample Size | Effective Sample Size | Nominal  Type I Error Rate | GEE  Empirical  Type I Error Rate | Mixed Model Empirical  Type I Error Rate |
| 20 | 16 | 0.05 | 0.38 | 0.11 |
| 20 | 16 | 0.05 | 0.36 | 0.10 |
| 20 | 18 | 0.05 | 0.33 | 0.08 |
| 20 | 18 | 0.05 | 0.32 | 0.07 |
| 20 | 20 | 0.05 | 0.19 | 0.06 |
| 40 | 32 | 0.05 | 0.21 | 0.06 |
| 40 | 33 | 0.05 | 0.20 | 0.06 |
| 40 | 35 | 0.05 | 0.17 | 0.05 |
| 40 | 36 | 0.05 | 0.16 | 0.06 |
| 40 | 40 | 0.05 | 0.11 | 0.05 |
| 80 | 64 | 0.05 | 0.11 | 0.06 |
| 80 | 66 | 0.05 | 0.10 | 0.05 |
| 80 | 70 | 0.05 | 0.09 | 0.05 |
| 80 | 72 | 0.05 | 0.09 | 0.05 |
| 80 | 80 | 0.05 | 0.08 | 0.05 |
| 160 | 128 | 0.05 | 0.07 | 0.05 |
| 160 | 131 | 0.05 | 0.07 | 0.05 |
| 160 | 141 | 0.05 | 0.06 | 0.05 |
| 160 | 144 | 0.05 | 0.07 | 0.05 |
| 160 | 160 | 0.05 | 0.06 | 0.05 |
| 200 | 160 | 0.05 | 0.07 | 0.05 |
| 200 | 164 | 0.05 | 0.06 | 0.05 |
| 200 | 176 | 0.05 | 0.07 | 0.05 |
| 200 | 180 | 0.05 | 0.06 | 0.05 |
| 200 | 200 | 0.05 | 0.06 | 0.05 |